10

Language Documentation & Conservation Special Publication No. 15
Reflections on Language Documentation 20 Years after Himmelmann 1998
ed. by Bradley McDonnell, Andrea L. Berez-Kroeker & Gary Holton, pp. 100–109
http://nflrc.hawaii.edu/ldc/
http://hdl.handle.net/10125/24812

Reflections on funding to support documentary linguistics

Gary Holton
University of Hawai'i at Mānoa

Mandana Seyfeddinipur
Endangered Languages Documentation Programme

Funding for documentary linguistics has changed dramatically over the past two decades, largely due to the emergence of dedicated funding regimes focused on endangered languages. These new regimes have helped to shape and reify the field of documentary linguistics by facilitating and enforcing best practices and integrating archiving into the documentation process. As a result both the pace and quality of documentation have improved dramatically. However, several challenges remain, and additional efforts are needed to ensure the sustainability of funding for language documentation efforts. In particular, more funding needs to be allocated toward training and capacity building in under-resourced regions.

1. Dedicated funding regimes for documentary linguistics The development of documentary linguistics over the past two decades is inextricably tied to the development of dedicated funding regimes which support the collection and organization of language documentation records. Best practices have been codified in grant proposal requirements, and proposal guidelines have likewise been shaped by emerging best practices in language documentation. Perhaps the greatest effect of these dedicated funding regimes has been to increase the valorization of the products of documentation and encourage or even enforce the archiving of those products. While there is great need for linguists to devote more time to language documentation activities, the academic reward system continues to place greater value on publications than on archival collections generated by language documentation activities (cf. Berez-Kroeker et al. 2018). Dedicated funding regimes provide a countermeasure to the established reward systems by incentivizing documentary activities which would not otherwise be highly valued within the academy.

The best-known of these are the large privately-funded schemes such as the Endangered Languages Documentation Programme (ELDP) and the Documentation of

ISBN: 978-0-9973295-3-7

Endangered Languages Program (DOBES) and government-funded regimes such as the United States National Science Foundation Documenting Endangered Languages (NSF-DEL) initiative, but there are many smaller private and public funding regimes as well, such as the Foundation for Endangered Language and the Endangered Languages Fund.[1] A detailed review and typology of the myriad funding regimes is beyond the scope of this chapter. Instead, we focus here on the impact that these programs have had on documentary linguistics, particularly with respect to documentation practices, archiving, and triage. We conclude with a discussion of the sustainability of funding for documentary linguistics.

2. Enforcement of best practices The emergence of dedicated funding regimes has both enforced and facilitated best practices in language documentation. At one time it was common practice for linguists to make use of professional quality recording equipment, such as the Nagra III-NP open reel recorder, released in 1958. This equipment could cost the equivalent of $10,000 or more in todays dollars and produced high-quality recordings with a relatively long shelf life.[2] However, by the late 20th century most linguists were making recordings using inexpensive consumer-grade cassette recorders, with little attention to long-term preservation. The digital revolution at the turn of the 21st century provided an impetus for change which was reinforced by the requirements of funders. Dedicated funding regimes proved not only willing to fund higher quality and more expensive recording equipment; they required the use of such equipment, including semi-professional digital video cameras and audio recorders and high quality external microphones. Moreover, not only do these funding regimes provide funding for such equipment, they also provide for or encourage training in the proper use of such equipment. For example, ELDP runs annual trainings for new grantees of ELDP-funded projects. Though it doesn't directly provide such training, the NSF-DEL program regularly funds training workshops such as the Institute on Collaborative Language Documentation (CoLang). Participation in these workshops generally increases the chances of success for an NSF-DEL grant application. Taken together, these efforts have helped to cement an accepted best practice for documentary linguistics.

3. Archiving requirement One of the most tangible effects of the dedicated funding schemes is the enforcement of archiving requirements. While archiving has always been a key part of the language documentation process, it has often been considered to be secondary to the production of descriptive materials such as grammars, dictionaries, and text collections. In the early days of documentary linguistics—prior to the Chomskyian turn—primary materials typically found their way to archives only upon the death of the collector, not upon the completion of the project. New funding regimes now require that archiving is completed as an integral part of the project. In fact, some schemes, including ELDP, view archiving as the primary goal of the funding program. The DOBES and ELDP schemes created digital archives from scratch for the express purpose of housing materials collected by their grantees. Arguably, one of the major contributions of the DOBES and ELDP schemes has been to develop and promulgate an integrated model of language

[1]For more information on these funding regimes see the relevant websites: `eldp.net`, `dobes.mpi.nl`, `www.nsf.gov/funding/pgm_summ.jsp?pims_id=2816`, `www.ogmios.org`, `www.endangeredlanguagefund.org`.

[2]For example, the Nagra 4SJ retailed for 11,539 Deutschmarks in 1972, the equivalent of over US$21,623 in 2018 when adjusted for inflation.

documentation which views archiving as the primary focus and primary outcome of the documentation process. Funding schemes enforce this model by requiring archiving as a primary output and by requiring grantees to develop an archiving plan as part of their project proposals. In the 11 years of its existence the DOBES programme funded 67 projects, most of which resulted in large, multimedia digital archive collections. Since its inception in 2002 ELDP has funded more than 400 projects of varying sizes, most resulting in a digital deposit in the Endangered Language Archive (ELAR). ELDP makes clear the priority placed on archiving. Since 2012 ELDP has enforced progressive depositing by requiring grantees to archive recordings and annotations annually annually and making disbursement of the next tranche of funding contingent on depositing.

Other funders have followed suit by strongly encouraging archiving as a project outcome. For example, the NSF-DEL program now requires that all applicants include an archiving plan which includes a letter of support from a repository which has agreed to accept the applicant's deposit. In addition, NSF-DEL applications must list the location of archiving in the application summary and discuss "plans and methodology for the sustainable, long-term archiving of all data" in the project proposal (National Science Foundation 2016). Perhaps most significantly from the point of view of funding, all of these funders now recognize archiving as a legitimate expense of language documentation work and allow for the costing of archiving in project budgets. ELDP achieves this by providing dedicated archiving facilities for grantees (through ELAR); NSF-DEL achieves this by allowing archiving to be costed in project budgets. Both approaches recognize that accessioning and sustaining digital language deposits is not without cost. In cases where the funder has not provided a dedicated repository, archiving with an external language archive can consume as much as 8% of project costs (DELAMAN 2014).

We like to think that the motivations for language documentation are altruistic and that linguists create archival deposits because they see value in an integrated model of language documentation and archiving. But given the competing demands on a researcher's time, the archiving incentives and requirements imposed by dedicated funding regimes have provided the impetus for the development of hundreds of new language documentation collections. It should be noted that other types of motivation may also help to incentivize language archiving as well. For example, the University of Hawai'i at Mānoa recently implemented an archiving requirement as part of its graduate program in linguistics (Berez-Kroeker 2015). All PhD candidates are required to submit proof of deposit in writing from the archive director to the dissertation committee before the dissertation can be approved. It is probably too early to tell whether such academic requirements will have the same force as financial incentives imposed by granting agencies.

4. Triage New funding regimes for endangered language documentation have also contributed to the development of a more sophisticated notion of triage for language endangerment. Given limited and finite resources, funding agencies must consider a number of different factors in order to set priorities for which language documentation projects should be funded. Beyond the obvious factor of whether the applicant is actually qualified and has the capacity to complete a proposed project, and whether there is evidence of community participation in the project, funders must weigh factors relating to the urgency of the documentation project. Degree of language endangerment is obviously an important consideration in this decision, but many other factors play a role as well.

Obviously we must prioritize documentation of a language whose last speakers may pass away within the next few years over the documentation of a language with millions of speakers, but often the choice is not so clear. For example, a highly endangered language from a very well-documented family may warrant lower priority for documentation than a relatively viable language isolate with no prior documentation. Documenting the world's linguistic diversity requires an approach which is also informed by our knowledge of genealogical relationships. Similarly, a language with unique and previously undocumented typological features may deserve higher priority for documentation than one with more typical features. Sometimes priorities for language documentation may be based on the contexts which are available for documentation. A relatively vital language may exhibit specialized domains of language use and language knowledge which have been lost in more endangered languages and thus are no longer available for documentation. In such cases the documentation of ethnobotanical knowledge or ritual language, for example, in a less endangered language may be prioritized over the documentation of prosaic language in a more endangered language. Of course, the number of factors involved means there is no single right way to prioritize language documentation efforts, but the need to effectively allocate funding has brought the issue of triage to fore of the language documentation enterprise.

5. Challenges As discussed above, the changes in the funding landscape over the past two decades have on the whole been beneficial for the field of language documentation, facilitating best practices and recognizing archiving as a fundamental part of the documentation process. But the emergence of large, dedicated funding regimes has also brought some challenges. Perhaps the greatest among these is the reinforcement of a distinction between documentation and revitalization/reclamation. Most funders of documentary projects prioritize documentation, often to the exclusion of reclamation efforts. For example, ELDP does not fund revitalization or language maintenance projects, taking a back seat in what can easily be seen as another colonial intervention. The justification for this focus on documentation is understandable, given the urgent need to create a record of languages while they are still spoken. However, this approach has both theoretical and practical limitations. Many language communities view documentation and reclamation efforts as intrinsically related and may even view documentation as a means to language reclamation rather than an end unto itself (cf. Fitzgerald 2017). This is especially true in the North American and Australian contexts. A single-minded focus on documentation is thus antithetical to some indigenous conceptualizations of language. Moreover—and in part as a consequence—in practical terms it is often impossible to separate documentation and reclamation efforts. Is a speaker making a recording in order to create a documentary record of her language or in order that her grandchildren may learn her language? Often there is no clear answer.

Documenters can exploit the fuzziness of this distinction in order to engage in reclamation efforts, but the perception of an artificial distinction between documentation and reclamation remains problematic. On the other hand, the situation is different for sub-Saharan Africa where language reclamation is not inextricably linked to documentation. In other words, community and speakers are happy to engage in documentation but may have no interest in reclamation as the language ecology, the pervasive multilingualism on the ground is the basis for a different language conceptualization (McGill & Austin 2012, Seyfeddinipur 2016, Seyfeddinipur & Chambers 2016).

Documenters can exploit the fuzziness of this distinction in order to engage in reclamation efforts, but the perception of an artificial distinction between documentation and reclamation remains problematic.

Another challenge for new funding schemes is the difficulty of striking a balance between enforcing accepted best practices and encouraging innovation. Given the urgency of endangered language documentation, funders are understandably hesitant to support high-risk projects which employ unproven methodologies. However, enforcement of best practices can go too far, leading to a potential stifling of innovation, as researchers adapt their work to fit the funding requirements rather than innovating new approaches. Funders may be reluctant to support new methods, such as "respeaking" techniques (cf. Reiman 2010), because they fall outside the boundaries of accepted practice. Yet the field desperately needs new, innovative methodologies in order to accelerate progress in language documentation, and developing these new methods may require support for unorthodox and untried approaches.

Another challenge is changing old habits and restrictive conceptualizations of language as a purely auditory phenomenon. Establishing best practices in the use of video in documentation remains a challenge despite the fact that there is thorough theoretical and methodological grounding for language a fundamentally multimodal phenomenon (Floyd 2016, Dingemanse 2013, Seyfeddinipur 2012). The value of video to both the scientific record and the community and their descendant is clear. Still, a casual review of deposits in DELAMAN archives reveals a strong reliance on audio, including many audio-only documentations. (This is true even in cases where there are no community restrictions on video use.) It is important to note that this is not a technological problem about how to use a video camera. Video equipment is now low-cost, portable and easy to use, and training in the use of video is readily available. Rather, the problem lies is convincing documenters that these non-auditory features of language are worthy of documentation. This constitutes a major challenge for funders trying to support high quality projects which do create a multipurpose (and multimodal) record.

One of the ironies of documentary linguistics is that linguists in regions which have the greatest threat to linguistic diversity tend to have the least access to funding for endangered language documentation. This exclusion is explicit in many national funding regimes. For example, the US National Science Foundation allocates funding only to scholars based at US institutions. And the DOBES programme required that the project leader have an affiliation with a German host institution. ELDP has been more effective than other regimes by having no restrictions on host institution locations, thus allowing funding world wide. Moreover, ELDP offers Small Grants which have no restrictions on academic qualification, allowing non-academics to apply. Nevertheless, more than 75% percent of ELDP funding has been allocated to institutions in just five countries (see figure 1).

That is not to say that all funding is going to documentation of languages in the US, Europe and Australia; merely that ELDP grants tend to be hosted by institutions in those countries. Some of this funding may actually go to scholars from outside those regions who are studying or pursuing postdoctoral research based at US, European or Australian institutions. Moreover, ELDP has made significant efforts to provide training for scholars outside US, Europe and Australia in order to encourage more competitive applications from those regions regions and to carry the documentation agenda into the areas where the documented languages are spoken. And national institutions outside the US, Europe and Australia are beginning to prioritize language documentation, as

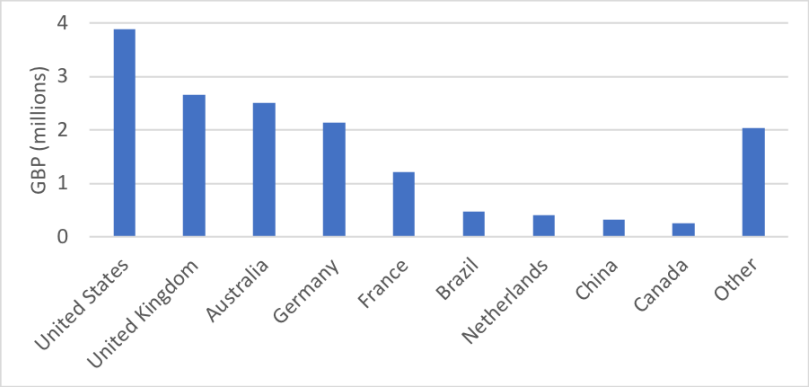

Figure 1: ELDP funding by location of host institution (2003-2016)

evidenced for example by the explicit reference to endangered languages in the 2016 call for proposals by the Indonesian Science Fund.[3] Nevertheless, funding for language documentation by scholars without an affiliation in the US, Europe and Australian remains a challenge—a point to which we return below.

6. Sustainability of funding for documentary linguistics Sustainability of funding for language documentation remains an ongoing challenge. While several large funding initiatives have emerged in the past two decades, some of these have already been discontinued. And while these funding programs have accomplished a lot, much more remains to be done. DOBES—once the largest funder of language documentation projects—ended its program in 2011. Throughout the 11 years of its existence the DOBES program was able to fund the creation of archival language documentation corpora for just 67 languages, representing less than one percent of the world's languages. While most linguists view language documentation as an ongoing effort, private funders typically view the individual projects they fund as finite, with a clearly defined end points.[4] Government funding bodies face similar constraints, as their constituents would like to see the task of documentation completed. The NSF Documenting Endangered Languages program is regularly threatened by legislative budget cutting. To a certain extent these sustainability issues are a direct result of the way the endangered languages "crisis" has been marketed to potential funders, both private and public. As a crisis, the endangered languages problem should be solvable and time-limited. That is, in characterizing the endangered languages issue as a crisis we have inadvertently defined it as a problem of finite proportions. To a certain extent this is of course true, but in practical terms the documentation of the world's endangered languages is not likely to be completed quickly.

Even if we could consider a typical 3-year major documentation project to be sufficient for the creation of a documentary collection, current funding regimes are now funding just a handful of such projects per year, perhaps at most 10-20. At that rate it will

[3]Dana Ilmu Pengetahuan Indonesia. `http://www.dipi.id`

[4]While individual projects may be of defined duration, the funding schemes themselves may be ongoing. Arcadia, the private donor of ELDP just renewed the funding in 2014 with another £7.2 million for five more years as they clearly saw that ELDP was only able to scratch the surface in the twelve years since the programme's inception.

take at least 150 years to document the slightly less than half of the world's languages which are today considered to be endangered. And during that time even more languages will become endangered, necessitating even more documentation projects. But a typical 3-year documentation project should be considered the bare minimum for creating a record of a language. It is much more common for linguists to devote entire careers to documenting a language. Seen in this light the endangered language documentation "crisis" is not likely to end soon but will rather be an ongoing effort which will need to be carried out over generations. It is unlikely that dedicated funding regimes—whether public or private—will survive over such a long time scale. Hence, linguists will need to look elsewhere to fund language documentation work in the future.

One obvious possibility is to return to the sources which have traditionally funded linguistic research, namely, national funding bodies such as the National Science Foundation (NSF) in the United States or the Deutsche Forschungsgemeinschaft (DFG) in Germany. This approach would have the effect of reintegrating language documentation back into mainstream linguistics. This approach comes with some risk as well. First, the loss of dedicated funding regimes would force language documentation proposals to compete directly with projects in theoretical linguists, ostensibly leading to less funding for language documentation and to the watering-down of language documentation efforts as applicants introduce theoretical components to their proposals in order to attract funding. Second, the loss of dedicated funding regimes could lead to less enforcement of best practices and less emphasis on archiving. This is especially problematic as many linguistics curricula have not been updated to train the new generations in documentary linguistics theory and methods. As discussed above, dedicated funding has been the primary factor driving the creation of digital multimedia collections as archival deposits; it is not clear that the more general funding schemes would maintain the same strict archiving requirements implemented by EDLP, DOBES, and NSF-DEL. Third, the loss of dedicated funding regimes could lead to less reliance on a pool of expert reviewers who are able to assess the priorities for documentation and the ability of applicants to successfully complete documentation projects. Finally, reliance on more general linguistics funding schemes would almost certainly result in a substantial reduction of funding for endangered language documentation.

A more effective solution to the sustainability problem may be to focus efforts on training local students and scholars, as ELDP has done since 2009. Local linguists will likely have better access to and understanding of the area and the communities. They themselves and their students may be speakers themselves who have a strong interest in documentation and one would hope for a domino effect of training the trainers who in turn train the new generation of documentary linguists working on their own languages. This would also address the issue of not having enough documentary linguists stemming the tide against language loss.

Another approach is to train native speakers to document their own languages. This approach has been advocated by several authors and has motivated the development of regular training institutes for North American languages, such as the American Indian Language Development Institute, the Institute for Collaborative Language Documentation, and the Canadian Indigenous Languages and Literacy Development Institute (Genetti & Siemens 2013, Florey & Himmelmann 2008). Funding training efforts has a number of advantages over funding linguists to do documentation work. The most obvious advantages are economic: native speakers generally do not need to travel in order to engage in documentation activities, so they can commit much greater amounts of time

to the work than can an outside linguist, at a much lower cost. However, even more significant are the quality advantages. Native speakers bring meta-linguistic knowledge which is often inaccessible to outside linguists or else only gained after years of study. Fundamental language documentation tasks such as transcription and annotation are much more effectively completed by native speakers than by outside linguists. Allocating additional funding to training a new generation of language documenters in under-resourced regions will do much more to address the endangered language crisis than will funding allocated to the current generation of documenters in the Global North.

However, training takes time. Developing new capacity in language documentation across the Global South and other under-resourced areas will require a significant commitment, both in terms of time from linguists and in terms of money from funding agencies. It affords hub building to ensure sustained support systems for the communities who are documenting their languages but the return is a major step in the right direction in safeguarding endangered languages. We would do well to turn our attention to language documentation training efforts while funders are still attuned to the endangered language crisis. Once these programs cease, it will be even more difficult to develop this new capacity.

References

Berez-Kroeker, Andrea. 2015. Reproducible research in descriptive linguistics: Integrating archiving and citation into the postgraduate curriculum at the University of Hawai'i at Mānoa. In Amanda Harris, Linda Barwick & Nick Thieberger (eds.), *Research, Records and Responsibility: Ten years of PARADISEC*, 39–51. Sydney: University of Sydney Press. http://hdl.handle.net/2123/16674.

Berez-Kroeker, Andrea L., Lauren Gawne, Susan Kung, Barbara F. Kelly, Tyler Heston, Gary Holton, Peter Pulsifer, David Beaver, Shobhana Chelliah, Stanley Dubinsky, Richard Meier, Nicholas Thieberger, Keren Rice & Anthony Woodbury. 2018. Reproducible research in linguistics: A position statement on data citation and attribution in our field. *Linguistics* 57(1). 1–18. doi:10.1515/ling-2017-0032.

DELAMAN. 2014. Report of the DELAMAN costing case study. Accessed 2018-05-11. http://hdl.handle.net/11122/6928.

Dingemanse, M. 2013. Ideophones and gesture in everyday speech. *Gesture* 13. 143–165. doi:10.1075/gest.13.2.02d.

Fitzgerald, Colleen. 2017. Understanding language vitality and reclamation as resilience: A framework for language endangerment and 'loss' (Commentary on Mufwene). *Language* 93(4). e280–e97.

Florey, Margaret & Nikolaus P. Himmelmann. 2008. New directions in field linguistics: Training strategies for language documentation in Indonesia. In Margaret Florey (ed.), *The endangered languages of Austronesia*, 121–140. Oxford: Oxford University Press.

Floyd, Simeon. 2016. Modally hybrid grammar? Celestial pointing for time-of-day reference in Nheengatú. *Language* 92(1). 31–64. doi:10.1353/lan.2016.0013.

Genetti, Carol & Rebecca Siemens. 2013. Training as empowering social action: An ethical response to language endangerment. In Elena Mihas, Bernard Perley, Gabriel Rei-Doval & Kathleen Wheatley (eds.), *Responses to language endangerment: In honor of Mickey Noonan*, 59–77. Amsterdam: John Benjamins.

McGill, Stuart & Peter K. Austin. 2012. Introduction and list of contributors (LDD 11). In Stuart McGill & Peter K. Austin (eds.), *Language Documentation and Description, vol 11.*, 5–27. http://www.elpublishing.org/PID/126.

National Science Foundation. 2016. Documenting Endangered Languages (DEL). NSF 16-576. Accessed 2018-05-11. https://www.nsf.gov/pubs/2016/nsf16576/nsf16576.htm.

Reiman, D. Will. 2010. Basic oral language documentation. *Language Documentation & Conservation* 4. 254–268.

Seyfeddinipur, Mandana. 2012. Reasons for documenting gestures and suggestions for how to go about it. In Nicholas Thierberger (ed.), *The Oxford handbook of linguistic fieldwork*, 147–165. Oxford: Oxford University Press.

Seyfeddinipur, Mandana (ed.). 2016. *African language documentation: new data, methods and approaches*. Honolulu: University of Hawai'i Press. http://hdl.handle.net/10125/24647.

Seyfeddinipur, Mandana & Mary Chambers. 2016. Language documentation in Africa: Turning tables. In Mandana Seyfeddinipur (ed.), *African language documentation: new data, methods and approaches*, 1–7. Honolulu: University of Hawai'i Press. http://hdl.handle.net/10125/24648.

Gary Holton
holton@hawaii.edu
orcid.org/0000-0002-9346-1572

Mandana Seyfeddinipur
ms123@soas.ac.uk